Estimating Forest Variables from Fusion of SAR and TM Data and Analytical Scattering and Reflectance Models

Mahta Moghaddam[1] and Jennifer L. Dungan[2]

[1]Jet Propulsion Laboratory, MS 300-227, 4800 Oak Grove Drive, Pasadena, CA 91109 USA
Tel.: (+1) 818-354-1591, FAX: (+1) 818-393-5285, Email: mahta.moghaddam@jpl.nasa.gov

[2]CSUMB, NASA Ames Research Center, MS 242-4 Moffett Field, CA 94035-1000 USA
Tel.: (+1) 650-604-3618, FAX: (+1) 650-604-4680, Email: jdungan@mail.arc.nasa.gov

ABSTRACT

A method for simultaneous integration of synthetic aperture radar (SAR) and optical remote sensing data in an estimation algorithm is presented which results in estimates of foliage mass over a larger range of values and more accurately than would be possible with either data type alone. The improved estimates are expected to result in more accurate calculation of ecosystem exchange from biogeochemistry models. The solution uses simplified closed-form models of scattering and reflectance derived from more complicated numerical models in a nonlinear estimation algorithm. Results are compared with available field measurements for 25 reference plots. A thorough error analysis is carried out to characterize the statistical accuracy of the estimation results with respect to errors in the microwave scattering and optical reflectance models. The foliage mass data are ultimately to be used to derive leaf area index (LAI), an essential driving variable for forest process models.

1. INTRODUCTION

Recent research and resulting achievements in single sensor analysis, along with the more widespread availability of several types of remote sensing data, create the possibility for combining those data and analyses into a single algorithm to invert for, or estimate, their key variables. Microwave and optical range frequencies are sensitive to variables defining vegetation canopies and contain independent information. Hence, their combination in a single algorithm should improve the resulting estimates by reducing uncertainty about estimates made from such imagery, as well as creating the possibility to extend the range of validity of the variable estimates. The latter is due to the sensitivity of each data type to different ranges of stand variables.

We have previously demonstrated these advantages through a regression-type analysis using field data of leaf biomass and near-simultaneous AIRSAR and Landsat TM data in the H.J. Andrews forest in Oregon [1]. Here, we expand this work by replacing the regression analaysis for radar data with a scattering model-based solution, and that for Landsat TM with a reflectance model-based solution. The numerical scattering model is used to simulate all channels of radar data assuming the knowledge of forest variables over a "typical" range and to express the generated results as closed-form functions, such as polynomials, of a small number of independent variables (in this case, 1). A similar approach is used with the numerical reflectance model based on a radiative transfer forlmulation [2] and incorporates a biophysically based leaf scattering coefficient with an absorption coefficient as determined by leaf thickness and the major biochemical constituents. The model can include variations in the canopy architecture through a variety of leaf angle distributions and soil background through soil reflectance. The resulting simple functions are used in a nonlinear optimization algorithm to estimate foliage mass. Allometric relations of the species under study will be used to relate the two variable sets needed in backscatter and reflectance models. Here the knowledge of species type will be assumed, and the variable set will be limited to one variable, foliage biomass.

Our multisensor variable estimation (or inversion) strategy can be summarized as optimizing the solution to a nonlinear system of equations. The system of equations is simply:

$$D = F(X), \quad (1)$$

Where,

- $D = (d_1, d_2, d_3, \ldots, d_N)^t$ is the vector of all remote sensing data of all types to be included for each spatial unit (pixel)
- $X = (x_1, x_2, \ldots x_M)^t$, with $M<N$, is the vector of all the variables to be estimated at that spatial unit
- $F(X)$ is an operator of dimension NxM that is a nonlinear functional of X, and each of its elements relates one or more of the data points to one or more of the variables.

The goal is to solve for X. In the case of estimating foliage mass only, dimension of X is simply 1 (one). We still include several measurements on the left-hand side of Equation (1) to improve solution accuracy and solve for X over a wider range since each of the elements of D may be sensitive to different ranges of X.

The system of equations would then be written as

$$(d_1, d_2, d_3, \ldots, d_N)^t = (F_1(x_1), F_2(x_1), F_3(x_1), \ldots, F_N(x_1))^t \quad (2)$$

with x_1 = foliage mass. Each functional on the right-hand side could relate its corresponding measurement to foliage mass through various analytical and numerical models. For TM data, these would be any of the canopy reflectance models. For SAR data, they would be the numerical forest scattering models. Due to the nonlinearity of each element of F with respect to x_1 the solution to this system has to be carried out as a nonlinear optimization. In addition, since both the variable x_1 and the measurements in D are stochastic in nature, appropriate modifications to the optimization algorithm need to be made through inclusion of their covariance operators. The optimization process is best carried out through an iterative algorithm such the conjugate gradient method or various preconditioned versions of it [3].

In the ideal, the models comprising F should be physically based, D is in physical units of surface reflectance or backscatter and the data available for X are precise field measurements taken at known (to an uncertainty) locations. Previously [1], we have solved Equation (2) when the right-hand side was derived from regression analyses. Here, we propose work towards the ideal by replacing the regression-derived components of F with those based on analytically based models, improving the retrieval of physical units from Landsat and radar imagery, and using additional field data for development and validation of the model.

2. AVAILABLE DATA

The data to be used in this study are collected over the H. J. Andrews Forest in Oregon. This area consists of various dense old-growth conifer stands, with biomass values ranging from less than 100 tons/hectare to over 1000 tons/hectare. The average altitude is about 950 m, with the lowest and highest points at about 600m and 1700m, respectively. The forest stand characteristics have also been extensively documented. In particular, foliage biomass and leaf-area index (LAI) values for approximately 30 reference stands have been reported [4]. These were used previously to construct the SAR and TM regression models, as well as to validate the unified estimation algorithm results. The remote sensing data consist of polarimetric C-, L-, and P-band radar data from the JPL airborne SAR (POLSAR/AIRSAR), the C-band single-polarization data from the JPL topographic SAR (TOPSAR), and the Thematic Mapper (TM) data from Landsat, all acquired in late April 1998. The range pixel spacing of the POLSAR is 3.3m for C- and L-bands and 6.6m for P-band. The TOPSAR pixel spacing is 10m, and the TM pixel size is 30m. Radiometric and polarimetric calibrations have been carried out on the POLSAR data. Due to pronounced topography, the radiometric calibration involves an added step to remove the effect of local slopes. The Landsat TM data were acquired under almost cloud-free conditions. All radar data are coregistered to the TM data using the PCI software (Figure 1) through geometric iteration processes.

Of the 15 AIRSAR, 2 TOPSAR, and 6 TM principally independent data channels, only a subset is actually useful and practically independent. Based on our previous studies [1], we have determined the following channels to be uniquely useful for our study: TM bands 1,2,4,7, TOPSAR C-VV correlation, POLSAR P-HH+P-HV, POLSAR C-HV+(C-HH or C-VV) or L-HV+(L-HH or L-VV).

3. ANALYTICALLY BASED SCATTERING AND REFLECTANCE MODELS

The relationships between foliage mass and Landsat TM bands and various channels of the AIRSAR data were derived by simulating them using numerical reflectance and forest scattering models, respectively, and inputting typical ranges of canopy variables. Allometric relations were used to relate foliage mass to reflectance and scattering properties such as expected tree heights, branch densities, and diameters. The simulation results were fitted to polynomials with foliage mass as the independent variable. The underlying assumption here is that the forest variables estimated from the radar scattering models can be related to those that could be estimated from optical remote sensing data such as Landsat TM. In other words, the measurements of various sensor types should be expressed in terms of a unified variable set, which is accomplished by using allometric relations.

4. ESTIMATION ALGORITHM

Once both SAR and TM data are expressed as closed-form functions, in this case polynomials, of the independent variable foliage mass, they can be used simultaneously in an estimation algorithm. Here, we have used a nonlinear optimization technique using an iterative conjugate gradient algorithm. Solutions are found within a few (10-20) iterations given a tolerable error condition. Statistical accuracy of the estimates are studied by superimposing polynomial coefficient noise with uniformly Gaussian distribution, repeating the estimation many times, and averaging the results.

5. RESULTS

The above algorithm is applied to the radar and optical data to estimate foliage biomass within the H.J. Andrews forest. The results are compared to those generated previously where all data were related to foliage biomass through regression models. It will be shown that the analytical scattering models produce

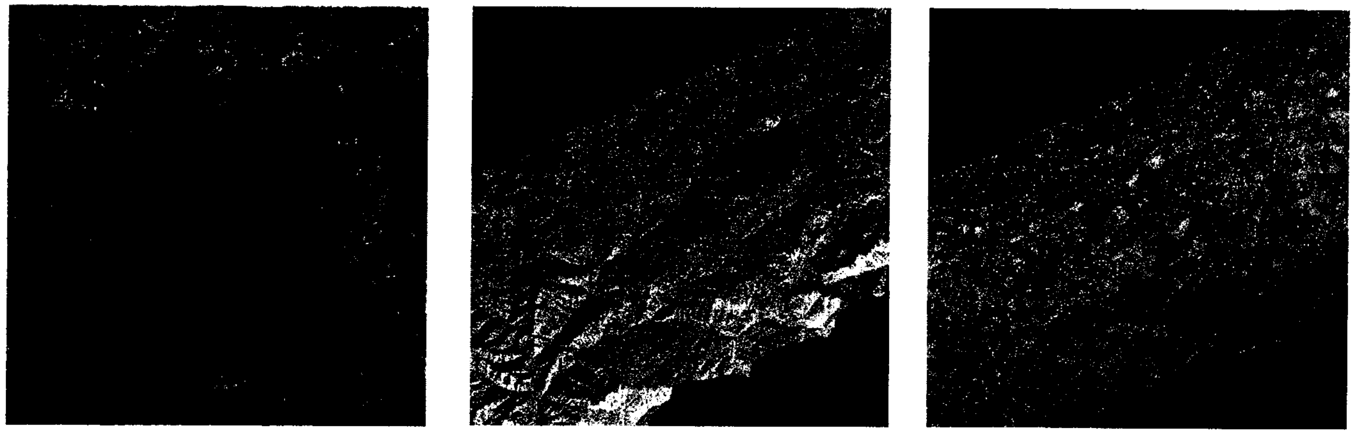


Figure 1. Coregistered Landsat TM (5,4,3), POLSAR, and TOPSAR data after geometric iteration and re-calibration.

more accurate estimation results that can be applied to a wider range of stands provided that accurate allometric relations are available. Numerical results will be shown at the presentation.

REFERENCES

[1] Moghaddam, M., J. Dungan, "Fusion of SAR and TM Data for Quantitative Estimation of Forest Variables Over an Extended Range of Validity," *Proc. IGARSS'00*, Honolulu, HI, 2000.

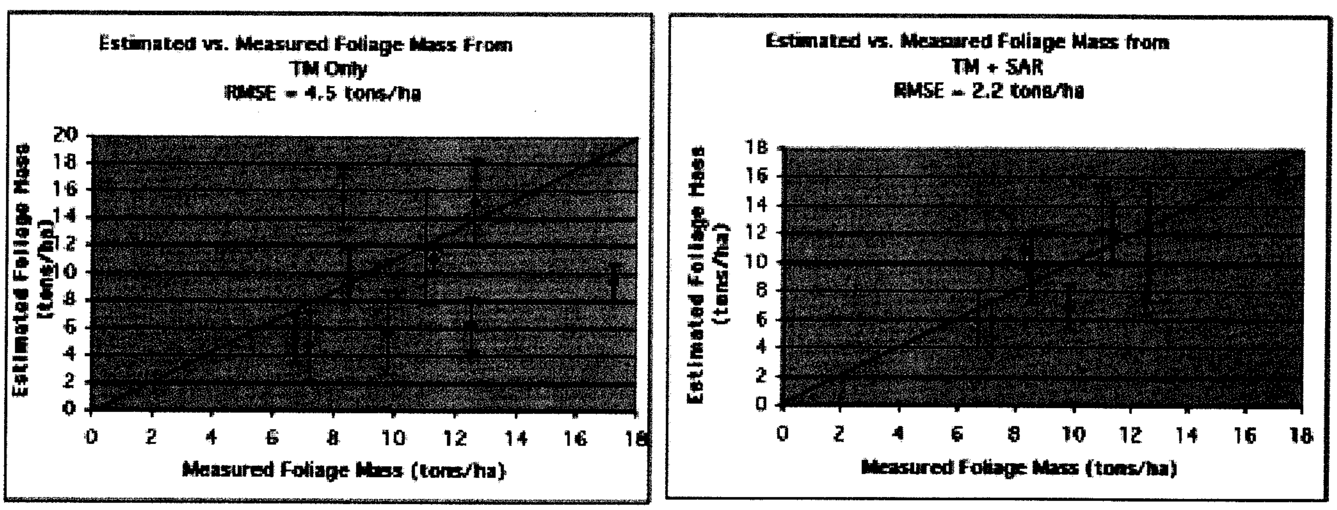
[2] Ganapol, B.D., L. F. Johnson, C. A. Hlavka, D. L. Peterson, and B. Bond, "LCM2: A coupled leaf/canopy radiative transfer model," *Remote Sens. Environ.*, 70:153-166, 1999.

[3] Moghaddam, M., and S. Saatchi, " Monitoring tree moisture using an estimation algorithm applied to SAR data from BOREAS," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, no. 2, pp. 901-916, 1999.

[4] Means, J. et al., "Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the western Cascades of Oregon," *Remote Sens. Environ.* 67:298-308, 1999.

ACKNOWLEDGMENT

This work was performed in part ~~by~~ at the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, and in part by the NASA Ames Research Center, Moffett Field, CA, both under contract from the National Aeronautics and Space Administration.



(a) (b)

Figure 2. Comparison of foliage mass estimation results with and without SAR data.